AFRL-HE-BR-TR-2006-0038

Air Force Research Laboratory

THE CHECK MARK PATTERN

Joel E. Michalek, PhD
Gary Henriksen, MD
Directed Energy Bioeffects Division
Human Effectiveness Directorate
Air Force Health Study Branch
2655 Flight Nurse Road
Brooks City-Base TX 78235

Pandu Kulkarni, PhD
Department of Mathematics and Statistics
University of South Alabama
Mobile, Alabama 36688

I. Jon Russell, PhD, MD
Department of Medicine
University of Texas Health Science Center at San Antonio
San Antonio, Texas 78284

Ram C. Tripathi, PhD
Department of Mathematics, Computer Science and Statistics
University of Texas at San Antonio
San Antonio, Texas 78249

Suojin Wang, PhD
Department of Statistics
Texas A&M University
College Station, Texas 77843

DIRECTED ENERGY BIOEFFECTS DIVISION
HUMAN EFFECTIVENESS DIRECTORATE
AIR FORCE HEALTH STUDY BRANCH
2655 FLIGHT NURSE ROAD
BROOKS CITY-BASE TX 78235

JULY 2006

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	
20-09-1995	Technical Report	

4. TITLE AND SUBTITLE The Check Mark Pattern

5a. CONTRACT NUMBER N/A

5b. GRANT NUMBER N/A

5c. PROGRAM ELEMENT NUMBER N/A

6. AUTHOR(S) Michalek, Joel E., AFRL/HEDA; Henriksen, Gary, AFRL/HEDA

Kulkarni, Pandu, Dept of Math and Statistics, Univ of South Alabama; Russell, I. Jon, Dept of Medicine, Univ of Texas Health Science Center at San Antonio
Tripathi, Ram C., Dept of Math, Comp Sc, and Statistics, Univ of Texas at San Antonio; Wang, Suojin, Dept of Statistics, Texas A&M Univ.

5d. PROJECT NUMBER 276700F1

5e. TASK NUMBER N/A

5f. WORK UNIT NUMBER N/A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

AFRL/HEDA
2655 Flight Nurse Road
Brooks City-
Base, Texas 7823-5137

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

None

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

AFRL-HE-BR-TR-2006-0038

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

None

14. ABSTRACT

We exhibit two studies (one epidemiological and one clinical), both with apparently paradoxical findings characterized by group (index versus control) similarity on the dependent (health) variable (Y) means, a significant group difference on the independent variable (X) means (index mean greater than the control mean) and a positive correlation between Y and X in the index group, causing index subjects with low values of X to have a lower Y mean than the controls and index subjects having high values of X to have a higher Y mean than the controls. This pattern has been called the "check mark" pattern. We predict this pattern using a linear model and use the model to estimate exposure effects in the epidemiologic study. Additionally, we show that a previously published study of the check mark pattern suggesting reverse causation is incorrect.

15. SUBJECT TERMS

Check mark pattern, Agent Orange, dioxin, Vietnam veterans, herbicides, insulin, taurine, tender point index fibromyalgia syndrome.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Julie N. Robinson, AFRL/HEDA
Unclassified	Unclassified	Unclassified	UU	29	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. Introduction page 3

2. Examples page 5

3. Modeling the Check Mark Pattern page 13

4. Discussion page 20

5. References page 24

6. Appendix page 28

TABLE OF TABLES

Table 1. Log(Dioxin+1) and Log(Insulin) by Group page 7

Table 2. Log(Dioxin+1) and Log(Insulin) by Dioxin Category page 8

Table 3. Tender Point Index and Taurine by Group page 11

Table 4. The Tender Point Index and Taurine by TPI Category page 12

Table 5. Category Definitions and Conditional Expectations page 13

TABLE OF FIGURES

Figure 1. Log(Insulin vs Log(Dioxin+1) by Group in the Air Force Health Study page 7

Figure 2. Log (Insulin) by Dioxin Category in the Air Force Health Study page 9

Figure 3.Taurine in CSF and the Tender Point Index in the Excitatory Amino Acid Study page 10

Figure 4. Taurine by TPI Category in the Excitatory Amino Acid Study page 12

ABSTRACT

We exhibit two studies (one epidemiological and one clinical), both with apparently paradoxical findings characterized by group (index versus control) similarity on the dependent (health) variable (Y) means, a significant group difference on the independent variable (X) means (index mean greater than the control mean) and a positive correlation between Y and X in the index group, causing index subjects with low values of X to have a lower Y mean than the controls and index subjects having high values of X to have a higher Y mean than the controls. This pattern has been called the "check mark" pattern. We predict this pattern using a linear model and use the model to estimate exposure effects in the epidemiologic study. Additionally, we show that a previously published study of the check mark pattern suggesting reverse causation is incorrect.

Key words: check mark pattern, epidemiologic studies, observational studies

1. INTRODUCTION

We consider medical studies of a health measure (Y) and an independent variable (X) in patients classified to index and control groups. In clinical studies the index subjects are generally "cases", who have been diagnosed as having a disease or syndrome and in epidemiological studies of health and exposure to some potentially harmful substance, the index subjects are identified as having been "exposed" to the substance. In the epidemiolgical study, X is a biomarker for exposure and in the clinical study X is a measure of disease or syndrome severity. In both types of studies we assume that the mean value of X in controls is less than the mean value of X in index subjects. Thus, if index and control subjects are otherwise similar one would expect that index subjects with low values of X should have a similar Y mean as control subjects and, if X is positively correlated with Y in the index group, one would expect that index subjects with large values of X would have a larger Y mean than controls. In fact, because index subjects have a larger mean value of X, one would further expect that the overall index Y mean would be greater than the overall control Y mean. In this paper, we consider studies where this expected pattern is not observed. The observed pattern is one in which the index and control means on Y are nearly equal and X and Y are positively correlated in the index group. In this situation, the Y mean among index subjects with low values of X will be less than the control Y mean and the Y mean among index subjects with high values of X will be greater than the control Y mean. Two examples of this "check mark" pattern are presented, modeled and interpreted. The first example arises in an epidemiological study and the second in a clinical study.

There is disagreement regarding the interpretation of the check mark pattern in epidemiological studies. A recent study of the pattern concluded that the pattern is suggestive of reverse causation, in which Y causes changes in X rather than X causing changes in Y (Flanders et al 1992). The Flanders paper has been cited by the Institute of Medicine (1994) in their recent review of Agent Orange research. We show that Flanders' arguments are incorrect and give alternative interpretations for the pattern.

The purpose of this paper is to describe the check mark pattern, both empirically (in Section 2) and with statistical models (in Section 3), to use the models to estimate exposure effects in the epidemiological example, to explore reverse order causation in the context of our models and to indicate an error in the reverse order causation argument of Flanders et al (1992). We do not intend to provide a statistical analysis of these data, as they have already been analyzed elsewhere in rigorous detail.

2. EXAMPLES

We present two examples of the check mark pattern in medical studies.

Example 2.1

The Air Force Health Study (AFHS) is a 20-year prospective epidemiological study of possible health effects from exposure to herbicides and their contaminant, 2,3,7,8-tetrachlorodibenzo-p-dioxin (dioxin) in veterans of Operation Ranch Hand, the unit responsible for spraying Agent Orange and other herbicides in Vietnam from 1962 to 1971 (Wolfe et al 1990). A group of Air Force veterans who served in Southeast Asia during the same period but who were not involved with spraying herbicides serve as controls. Control veterans were matched to Ranch Hand veterans on age, military occupation and race. The study compares the health, reproductive outcomes and mortality of Ranch Hand veterans with control veterans. Physical examinations were administered in 1982, 1985, 1987 and 1992. Examinations are planned for 1997 and 2002. Since 1987, a measurement of dioxin in serum has been used as the index of exposure (Wolfe et al 1992).

Nine hundred fifty two Ranch Hands and 1,281 controls attended the 1992 AFHS physical examination of whom 894 Ranch Hands and 1,084 controls had quantitatable dioxin results (Grubbs et al 1995). An investigation of endocrine function included serum insulin, measured in mIU/ml. Because we were interested in the

association between dioxin (X) and insulin (Y) in nondiabetic subjects, we excluded known diabetics and subjects with 2 hour post-prandial glucose levels greater than 200 mg/ml (125 controls and 115 Ranch Hands), leaving 959 controls and 779 Ranch Hands with insulin and dioxin levels for consideration.

We studied the relation between insulin and dioxin after both had been logarithmically transformed (the number 1 was added to the dioxin result before taking the logarithm) using linear regression and t-tests. The data are shown in Figure 1, with a least-squares line overlaid on the Ranch Hand data. The coefficient of log(dioxin+1) in the regression of log(insulin) in Ranch Hands is statistically significant (slope=0.18, standard error=0.031).

Figure 1: Log(Insulin) vs Log(Dioxin+1) by Group in the 1992 Air Force Health Study

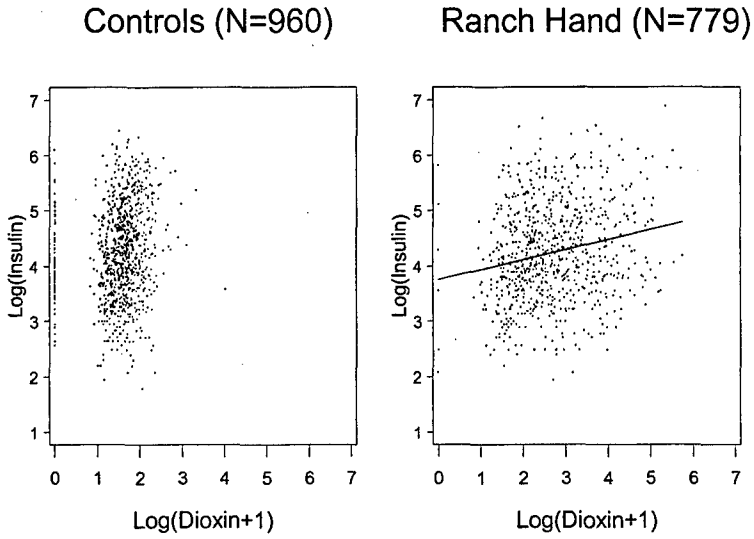


Table 1 gives summary statistics by group.

Table 1: Log(Dioxin+1) and Log(Insulin) by Group

Group	N	Log(dioxin+1) Mean (Std Dev)	Log(insulin) Mean (Std Dev)
Ranch Hand	779	2.7 (0.99)	4.24 (0.88)
Control	959	1.8 (0.44)	4.25 (0.83)

By design, the mean of log(dioxin+1) in Ranch Hands (2.7) is greater than the mean of log(dioxin+1) in controls (1.8). The group log(insulin) means were not significantly different [mean difference=-0.01, 95% CI: (-0.07, 0.09)].

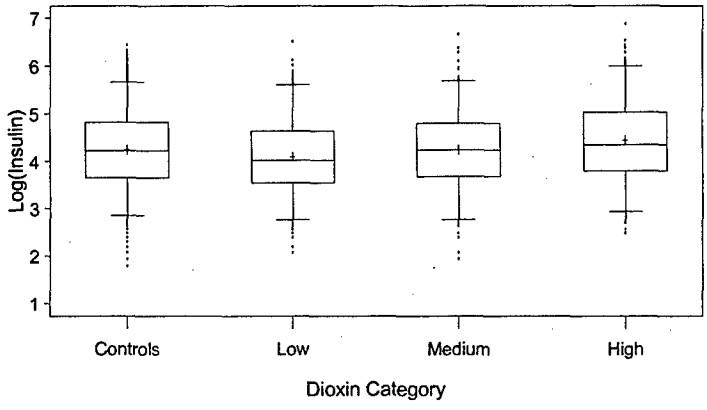
We stratified the data to four categories according to group and dioxin level. The first category was comprised of control subjects with dioxin less than 10 ppt, the value we regard as the threshold for background exposure. The remaining three categories, named "Low", "Medium" and "High" were determined by the cut points 10 ppt and 24 ppt in the dioxin distribution in the Ranch Hand (RH) group. Table 2 gives summary statistics for log(insulin) and log(dioxin+1) in each of these four dioxin categories.

Table 2: Log(Dioxin+1) and Log(Insulin) by Dioxin Category

Dioxin Category	**N**	**Log(dioxin+1) Mean (Std Dev)**	**Log(insulin) Mean (Std Dev)**
Control	942	1.5 (0.55)	4.2 (0.83)
RH Low	341	1.8 (0.44)	4.1 (0.84)
RH Medium	221	2.8 (0.24)	4.3 (0.87)
RH High	217	4.0 (0.59)	4.4 (0.90)

The data are summarized in Figure 2 by dioxin category, showing the check mark pattern.

Figure 2: Log(Insulin) by Dioxin Category in the Air Force Health Study



LEGEND: Low, Medium and High are Dioxin categories in Ranch Hand subjects. Whiskers extend to the 5th and 95th percentiles. Dots are lower and upper 5 percent. Rectangles are determined by quartiles, and the mean is indicated with a + .

Example 2.2

A recent study of excitatory amino acids in cerebral spinal fluid (CSF) and severity of pain or tenderness in patients with fibromyalgia syndrome (FS) included 41 patients diagnosed as having FS and 37 age, sex and ethnicity matched healthy normal controls; we call this study the CSF excitatory amino acid (CSF-EAA) study.

In the CSF-EAA study CSF samples were obtained from each individual by standard lumbar puncture. Taurine (Y) was measured in CSF as previously described using specific radioimmunoassays (Giovengo et al 1995) and was reported in ng/μl. The severity of tenderness (X) was measured by the tender point index (TPI), which involved digital palpation at 18 specific soft tissue sites (Wolfe et al 1990). In addition to

inquiring about pain induced by palpation, the physician closely observed the patient for any physical response and scored each site as follows: 0 (no pain), 1 (tenderness reported but without physical response), 2 ("semi-objective" tenderness demonstrated by a physical response such as a wince or withdrawal), 3 (very exaggerated physical response) and 4 (untouchable; patient will not allow palpation at a given site fearing unbearable pain). The diagnosis of FS was based on the continuous presence of musculoskeletal pain for 3 months and "semi-objective" tenderness at 11 or more of the 18 tender points (Wolfe et al 1990). The data are shown in Figure 3, with a least-squares line overlaid on the FS data. The coefficient of TPI in the regression of taurine in FS patients is statistically significant (slope=0.12, standard error=0.031).

Figure 3: Taurine in CSF and the Tender Point Index in the Excitatory Amino Acid Study

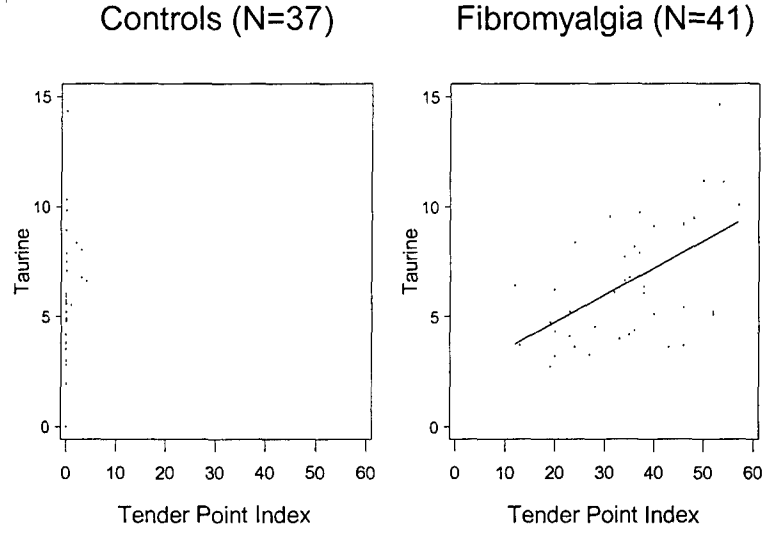


Table 3 gives summary statistics by group.

Table 3: Tender Point Index and Taurine by Group

Group	N	Tender Point Index Mean (Std Dev)	Taurine Mean (Std Dev)
Fibromyalgia Syndrome	41	34.9 (11.89)	6.6 (2.75)
Control	37	0.4 (0.98)	5.9 (2.61)

By design, the FS mean on TPI (34.9) is greater than the control mean (0.4). The mean value of taurine did not vary significantly with group [mean difference=1.13 ng/μl, 95% CI: (-0.06 ng/μl , 2.32 ng/μl)].

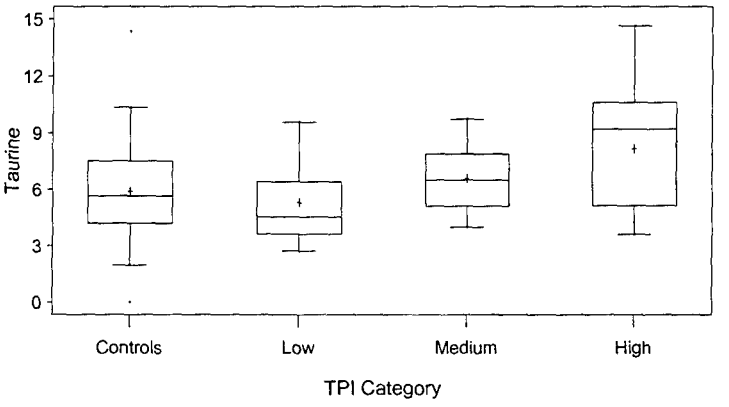
We stratified the data to four categories according to group and TPI level. The first category was comprised of control subjects. The remaining three categories, named "Low", "Medium" and "High" were determined by the tertiles (31 and 40) of the TPI distribution in FS patients. Table 4 gives summary statistics for taurine and the TPI in each of these four TPI categories.

Table 4: The Tender Point Index and Taurine by TPI Category

TPI Category	N	Tender Point Index Mean (Std Dev)	Taurine (ng/µl) Mean (Std Dev)
Control	37	0.35 (0.98)	5.88 (2.62)
FS Low	15	22.3 (5.62)	5.29 (2.26)
FS Medium	14	36.1 (2.43)	6.58 (1.80)
FS High	12	49.4 (4.21)	8.16 (3.49)

The data are summarized in Figure 4 by TPI category, showing the check mark pattern.

Figure 4: Taurine by TPI Category in the Excitatory Amino Acid Study



LEGEND: Low, Medium and High are TPI categories in FS patients. Whiskers extend to the 5th and 95th percentiles. Dots are lower and upper 5 percent. Rectangles are determined by quartiles and the mean is indicated with a +.

3. MODELING THE CHECK MARK PATTERN

3.1 Definition of the Check mark Pattern

The patterns in each of these two examples are caused by nearly equal group means on Y and a positive correlation between X and Y in the index group. The conditional means shown in Tables 2 and 4 then exhibit the pattern, with the control mean on the health variable (Y) being greater than the index mean in the Low category and an increasing sequence of conditional means on Y in the index categories (Low, Medium and High) determined by two cut points (c_{11} and c_{21}) in the distribution of the independent variable (X) in the index group, where c_{11} and c_{12} are approximate tertiles of X in the index group. The control category is comprised of control subjects, possibly truncated with X at or below a cut point c_{10}, which we assume as a high quantile of the distribution of X in the control group. To formalize this phenomenon, we introduce notation for the conditional expectations in Table 5, where group is indicated by j (control:j=0, Index:j=1).

Table 5: Category Definitions and Conditional Expectations

Category	**j**	**Interval**	**Label**	**$E(Y\|j,X\in A_k)$**
Control	0	$(-\infty, c_{10})$	A_0	μ_{y0}
Index Low	1	$(-\infty, c_{11})$	A_1	μ_{y1}
Index Medium	1	(c_{11},c_{21})	A_2	μ_{y2}
Index High	1	(c_{21},∞)	A_3	μ_{y3}

The check mark pattern is said to hold if $\mu_{y0}>\mu_{y1}$ and $\mu_{y1}<\mu_{y2}<\mu_{y3}$. In the AFHS, c_{10}=10 ppt, c_{11}=10 ppt and c_{21}=24 ppt and in the CSF-EAA study $c_{10}=\infty$, c_{11}=31 and c_{21}=40.

3.2 Statistical Models to Predict the Check mark Pattern

We propose standard linear statistical models to predict the observed patterns in Tables 2 and 4. Let δ_{ji} =1 if subject i is in group j, 0 otherwise, j=0,1, i=1, 2, ..., n_j, and consider the following linear model for subject i in group j,

$$Y_{ji} = \beta_0 + \beta_1[(X_{0i} - \mu_{x0})\delta_{0i} + (X_{1i} - \mu_{x1})\delta_{1i}] + \varepsilon_{ji}, \quad (1)$$

where, for subject i in group j, j=0,1, i=1,2, ..., n_j, Y_{ji} and X_{ji} are the observed values of Y and X, μ_{xj} is the population mean of X from group j and ε_{ji} is a random error independent of X_{ji} and δ_{ji} with mean 0 and variance σ^2. Under model (1), $E(Y_{0i})=E(Y_{1i})=\beta_0$, that is, the overall means on Y in the two groups are equal. Furthermore, if $\beta_1>0$, it is easy to verify that model (1) predicts the check mark pattern. Using model (1), we have

$$\mu_{y0}=\beta_0+\beta_1[E(X_{0i}|X_{0i}\leq c_{10})-E(X_{0i})]. \quad (2)$$

But, since c_{10} is a high quantile of the control distribution of X, the coefficient of β_1 in (2) is approximately zero. However,

$$\mu_{y1}=\beta_0+\beta_1[E(X_{1i}|X_{1i}\leq c_{11})-E(X_{1i})] \quad (3)$$

and the coefficient of β_1 in (3) is negative because c_{11} is an approximate lower tertile of the distribution of X in the index group. Thus, $\mu_{y0}>\mu_{y1}$. The remaining conditional means follow the ordering $\mu_{y1}<\mu_{y2}<\mu_{y3}$ because $\beta_1>0$.

In the special case that all of the controls have X=0 and the index subjects have positive values of X, as in the CSF-EAA study, we reduce model (1) to

$$Y_{ji} = \beta_0 + \beta_1[(X_{i1} - \mu_{x1})\delta_{1i}] + \varepsilon_{ji}. \qquad (4)$$

Once again, it is easy to verify that the reduced model (4) predicts equal means and the check mark pattern when $\beta_1>0$.

Model (1) implies that β_0 is the population mean of group 1 as well as of group 0. Therefore, β_0 is estimated by $\overline{Y}_0$, $\overline{Y}_1$ or $\overline{Y} = (n_0\overline{Y}_0 + n_1\overline{Y}_1)/(n_0 + n_1)$. Assuming a common variance in the two groups, it is better to estimate β_0 with $\overline{Y}$ because its variance is smaller than that of $\overline{Y}_0$ or $\overline{Y}_1$. Therefore, to accommodate this in our models, so that standard statistical software packages can be used, at the time of fitting the models, we replace the means μ_{xj} by their estimators $\overline{X}_j$, j=0,1. Based on (1), the conditional log(insulin) means are $\hat{\mu}_{y0} = 4.24$, $\hat{\mu}_{y1} = 4.08$, $\hat{\mu}_{y2} = 4.26$ and $\hat{\mu}_{y3} = 4.49$, which are very comparable to the observed means in Table 2. The predicted conditional means using model (4) for taurine, are $\hat{\mu}_{y0} = 6.25$, $\hat{\mu}_{y1} = 4.68$, $\hat{\mu}_{y2} = 6.37$ and $\hat{\mu}_{y3} = 8.03$, corresponding to the observed means in Table 4. Thus, model (1) and its reduction (4) describe the check mark patterns given in the two examples.

3.3 Reverse Causation

At this point we should mention that Flanders et al (1992) conducted a study of the check mark pattern using AFHS data. They studied two models; their first model considered a relationship between health and dioxin with health as the dependent variable and dioxin as the independent variable and their second model reversed the roles of these two variables. When analyzing their second model, they made the assumption that E(Y|j=0)=E(Y|j=1), i.e. the Ranch Hand and control health means are equal, while not making this assumption for their first model. As we show in the Appendix, under this assumption the coefficient of health in their second model is identically zero (which was not recognized by Flanders et al), making their model not very useful. This mathematical oversight caused them to believe that their second "reverse causation" model was the appropriate model and hence concluded that the check mark pattern suggested reverse causation. Therefore, the mathematical arguments used by Flanders et al to argue that the check mark pattern suggests reverse causation are incorrect.

To explore the possibility of reverse causation in the context of our model (1), suppose that reverse causation actually holds and consider a revision of model (1) with dioxin as the dependent variable and health as the independent variable, given by

$$Y^*_{ji} = \beta_0 + \beta_j[(X^*_{0i} - \mu^*_{x0})\,\delta_{0i} + (X^*_{1i} - \mu^*_{x1})\,\delta_{1i}] + \varepsilon_{ji}\,, \qquad (5)$$

where Y^*_{ji} is the dioxin level and X^*_{ji} is the health variable of subject i in group j, and μ^*_{xj} is the health variable mean in group j, j=0,1. Now assume, that the health variable means in the two groups are equal, E(X*|j=0)=E(X*|j=1). Then, using model (5) we find that $E(Y^*_{0i}) = E(Y^*_{1i})$, implying that the mean dioxin levels in the two group are the same,

which violates the known difference between the Ranch Hand and control dioxin distributions. Note, however, that this happens regardless of the assumption of equal group means. Hence, one cannot conclude reverse causation by a simple reversal of the roles of X and Y in model (1).

Now, one may argue that although model (5) does not predict the check mark pattern and so does not support reverse causation, that there might be another model that describes the relationship between Y* and X* and also predicts the check mark pattern. Hence, we consider a more general model

$$Y^*_{ji} = h(\Theta, X^*_{ji}) + \varepsilon_{ji}, \qquad (6)$$

where $h(\Theta, X^*_{ji})$ is some function of a vector of parameters Θ and X^*_{ji} and ε_{ji} has mean 0 and variance σ^2. Here, we are working under the assumption that Y^*_{ji} (dioxin) is the dependent variable and that X^*_{ji} (health) is the independent variable. Now, following Flanders argument, to produce the check mark pattern we need conditional expectations of (6) (given that dioxin is contained in an interval), such as

$$E(Y^*_{0i} | Y^*_{0i} \le c^*_{10}) = E[h(\Theta, X^*_{0i}) + \varepsilon_{0i} | Y^*_{0i} \le c^*_{10}]. \qquad (7)$$

However, the expectation on the right hand side of (7) violates the traditional definition of regression, as it involves the conditional expectation of some function of the independent variable conditional on the dependent variable. Therefore, the definition of our dependent and independent variables is in jeopardy. Hence, this model violates the traditional

approach to regression unless we reverse the roles of the variables, in which case model (1) appears to be appropriate. Therefore, we consider the statistical utility of model (1) in the Air Force Health Study.

3.4 Estimating Exposure Effects

Because the concept of "exposure" is central to the AFHS, one may define a subject as exposed (e=1) if his dioxin level is above c_{10} and unexposed (e=0) otherwise. Using model (1), the mean of Y among the exposed, μ_{yE}, can be written as

$$\begin{aligned}\mu_{y\bar{E}} &= E(Y_{ji}|e=1,\ j=0)P(j=0|e=1) + E(Y_{ji}|e=1,\ j=1)P(j=1|e=1)\\ &= \left\{\beta_0 + \beta_1[E(X_{0i}|X_{0i} > c_{10}) - \mu_{x0}]\right\}\frac{P(e=1|j=0)P(j=0)}{P(e=1)}\\ &\quad + \left\{\beta_0 + \beta_1[E(X_{1i}|X_{1i} > c_{10}) - \mu_{x1}]\right\}\frac{P(e=1|j=1)P(j=1)}{P(e=1)}.\end{aligned}$$

Similarly, the mean of Y among the unexposed is

$$\begin{aligned}\mu_{y\bar{E}} &= E(Y_{ji}|e=0,\ j=0)P(j=0|e=0) + E(Y_{ji}|e=0,\ j=1)P(j=1|e=0)\\ &= \left\{\beta_0 + \beta_1[E(X_{0i}|X_{0i} \le c_{10}) - \mu_{x0}]\right\}\frac{P(e=0|j=0)P(j=0)}{P(e=0)}\\ &\quad + \left\{\beta_0 + \beta_1[E(X_{1i}|X_{1i} \le c_{10}) - \mu_{x1}]\right\}\frac{P(e=0|j=1)P(j=1)}{P(e=0)}.\end{aligned}$$

However, the probabilities P(j=0) and P(j=1) are generally not estimable. Hence we consider the estimable difference of the Y mean of exposed index subjects and the Y mean of unexposed controls, given by $D = \mu_{yE1} - \mu_{y\bar{E}0}$, where $\mu_{yE1} = E(Y|e=1, j=1)$ and $\mu_{y\bar{E}0} = E(Y|e=0, j=0)$ which, under model (1), is given by

$$D = \beta_1[p_{\bar{E}1}(\mu_{xE1} - \mu_{x\bar{E}1}) + p_{E0}(\mu_{xE0} - \mu_{x\bar{E}0})],$$

where p_{Ek}=P(e=1|j=k), $p_{\bar{E}k} = P(e = 0 | j = k)$, $\mu_{xEk} = E(X | e = 1, j = k)$ and $\mu_{x\bar{E}k} = E(X | e = 0, j = k)$, k=0,1. Hence D is estimated by

$$\hat{D} = \hat{\beta}_1[\hat{p}_{\bar{E}1}(\overline{X}_{E1} - \overline{X}_{\bar{E}1}) + \hat{p}_{E0}(\overline{X}_{E0} - \overline{X}_{\bar{E}0})],$$

where $\overline{X}_{Ek}$ and $\overline{X}_{\bar{E}k}$ are the sample means of exposed and unexposed subjects in group k, k=0,1, $\hat{p}_{E1}$ is the sample proportion of exposed subjects in group 1 and $\hat{p}_{\bar{E}0}$ is the sample proportion of unexposed subjects in group 0. The estimate of the conditional standard deviation of $\hat{D}$, given X, is

$$\hat{\sigma}_{\hat{D}} = \hat{\sigma}\hat{\beta}_1[\hat{p}_{\bar{E}1}(\overline{X}_{E1} - \overline{X}_{\bar{E}1}) + \hat{p}_{E0}(\overline{X}_{E0} - \overline{X}_{\bar{E}0})].$$

Hence, to test the hypothesis H_o:D=0, we use the statistic $T=\hat{\beta}_1 / \hat{\sigma}\hat{\beta}_1$, distributed as t with n-2 degrees of freedom under H_o. In the special case that all of the index subjects are exposed and all of the controls are unexposed, D reduces to the difference of group means E(Y|j=1)-E(Y|j=0) and the appropriate test statistic is the ordinary two sample t-test.

For example, to estimate D in the Air Force Health Study data on log(insulin) and dioxin, we have $\hat{\beta}_1 = 0.18$, $\hat{\sigma}\hat{\beta}_1 = 0.031$, $p_{\bar{E}1} = 0.44$, $p_{E0} = 0.02$, $\overline{X}_{E1} = 3.384$, $\overline{X}_{\bar{E}1} = 1.836$, $\overline{X}_{E0} = 2.705$ and $\overline{X}_{\bar{E}0} = 1.501$. Therefore, $\hat{D} = 0.13$ and $\hat{\sigma}_{\hat{D}} = 0.022$. Hence a 95% confidence interval for D is 0.09 to 0.17 and the test statistic for H_o is T=5.81 (p<0.001).

4. DISCUSSION

There is currently no convincing explanation for the check mark pattern in the Air Force study. The relationship between dioxin and insulin could reflect direct causation, bias, reverse causation, or the effects of differential dioxin elimination. The association is consistent with direct causation through dioxin blocking insulin receptors. This hypothesis would explain the trend of increasing insulin with increasing dioxin in nondiabetic Ranch Hands and a decrease in mean insulin with increased dioxin in diabetic Ranch Hands (data not shown; see Grubbs et al 1995). However, this explanation is not adequate to also explain the near equality of the insulin means in Ranch Hands and controls. The near equality of observed group insulin means follows from standard bias models (Anderson et al 1980). If the cut point we regard as the threshold for exposure (10 ppt) truly separates exposed from unexposed, then 44% of the Ranch Hand group is unexposed, causing the observed group mean difference to be biased toward equality. This, together with a positive correlation between dioxin and insulin in Ranch Hand veterans, would produce the check mark pattern.

Reverse causation could occur through contaminated medications. It is possible that medications are contaminated with trace amounts of dioxin, as are many foods. If so, repeated doses of medications could, over a period of years, increase the dioxin body burden of subjects with disease. However, this possibility cannot explain the association between dioxin and log(insulin) because we restricted the analysis to nondiabetic subjects

who are not taking insulin-controlling medications. Nevertheless, there might be other sources of dioxin uptake which correlate with insulin that could account for the pattern.

The known association between dioxin half-life and percent body fat, a correlate of insulin, might also contribute to the association between insulin and dioxin. Obese persons have a longer dioxin half-life than lean persons (Michalek et al, 1995). Thus, veterans who are overweight are more likely to have elevated insulin levels, but are also more likely to retain their dioxin longer than lean veterans. However, the association between dioxin and log(insulin) remains significant even after adjustment for percent body fat.

Interpretation of the check mark pattern in the CSF-EAA study requires consideration of the mechanism of pain transmission (referred to as nociception). Nociception involves a complex series of electrochemical, receptor ligand interaction, and second messenger processes which sequentially occur in peripheral nerves, the spinal cord, and the brain. This process is further complicated because several steps are subject to down-regulation or "sensory interpretation" by inhibitory (antinociceptive) neurochemical, receptor ligand, and second messenger effects.

A very abbreviated view of the nociceptive process is that a painful peripheral stimulus initiates the electrochemical depolarization of a primary efferent neuron which releases substance P and excitatory amino acids (EAA) into the dorsal horn of the spinal cord (Malmberg and Yaksh 1992). Substance P facilitates the nociceptive activity of the EAA which induce the release of prostaglandins and, perhaps, nitric oxide from arginine.

The prostaglandin messenger then causes depolarization of the ascending spinal cord neuron which carries the nociceptive message to the brain.

Antinociception is accomplished though inhibitory contributions from a peptide fragment resulting from protease digestion of substance P (substance P_{1-7}), serotonin, taurine, and endogenous opioids (Hornfeldt et al 1992, Smullin et al 1990, Skilling et al 1990). When the spinal cord of rats was exposed to intact substance P, it caused the release of taurine, but substance P_{1-7} seemed to inhibit that release.

The actual role of taurine in nociception and antinociception is not yet clear. The tissue concentration of taurine decreases in the hypothalamus and lower brain-stem nuclei of rats experiencing acute pain (Palkovitz et al 1986). It should be noted, however, that the brain concentration of a given neurochemical may vary inversely with its concentration in the spinal cord (Sharma et al 1990). Intratheceally-injected taurine appeared to induce writhing pain in animals (Larson 1989). Conversely, taurine inhibited the pain-like responses of rats subjected to intrathecal injection of substance P or to the acetic acid-induced writhing test (Smullin et al 1990), and both of these effects could be blocked by administration of the taurine antagonist 6-aminoethyl-3-methyl-4H-1,2,4-bendo-thiadiazine-1,1-dioxide (TAG).

The correlation of taurine with the TPI measure of pain herein described as an example of the check mark pattern may simply represent the association of taurine with the nociception/antinociception process. In other words, the pattern might result from a combination of direct and reverse causation, with pain causing changes in taurine levels and taurine levels causing changes in pain. The FS patients were experiencing a range of

painful sensations prior to the TPI examination, while the normal controls were not experiencing pain de novo and the examination pressure was not perceived by them to be painful. It is possible that a number of other chemical mediators of nociceptive and antinociceptive processes would also exhibit a similar pattern if they are not too rapidly degraded in the course of normal homeostatic regulation.

We conclude that the check mark pattern might be caused by direct or reverse causation, misclassification bias, or differential dioxin elimination in the Air Force study and that the pattern may arise from a combination of direct and reverse causation in the CSF-EAA study. Thus, the reverse causation explanation for the checkmark pattern appears plausible, even though a previously published study of reverse causation is mathematically incorrrect. Finally, we have demonstrated that a simple linear model describes the pattern and that the model can be used to estimate exposure effects in the Air Force study.

5. REFERENCES

Anderson, S., Auquier, A., Hauck, W.W., Oakes, D., Vandele, W., Weisberg, H.I., Bryk, A.S. and Kleinman, J. (1980). Statistical methods for comparative studies. John Wiley and Sons, New York.

Flanders, W.D., Lin, L., Pirkle, J.L., and Caudill, S.P. (1992), "Assessing the direction of causality in cross-sectional studies," The American Journal of Epidemiology,135, 926-935.

Giovengo, S., Russell, I.J. and Larson, A.A. (1995). Amino acids in cerebral spinal fluid of patients with fibromyalgia. J Musculoskeletal Pain 3 (supple #1) 9,1995.

Grubbs, W.D., Lustik, M.B., Brockman, A.M., Henderson, S.C., Burnett, F.R., Land, R.G., Osborne, D.J., Rocconi, V.K., Schreiber, M.E., and Williams, D.E. (1995), The Air Force Health Study. An Epidemiologic Investigation of Health Effects in Air Force Personnel Following Exposure to Herbicides. 1992 Follow-up Examination Results, Armstrong Laboratory: Brooks Air Force Base, Texas.

Institute of Medicine (1994), Veterans and Agent Orange. Health Effects of Herbicides Used in Vietnam, Washington DC: National Academy Press.

Hornfeldt, C.S., Smullin, D.H., Schamber, C.D., Sun, X. and Larson, A.A. (1992). Antinociceptive effects of intrathecal taurine and calcium in the mouse. Life Sciences 50, 1925-1934.

Larson, A.A. (1989). Intrathecal GABA, glycine, taurine, or beta-alanine elicits dyskinetic movements in mice. Pharmacol Biochem Behav. 32, 505-509.

Malmberg, A., Yaksh, T. (1992). Hyperalgesia mediated by spinal glutamate or substance P receptor blocked spinal cyclooxygenase ihibition. Science 257, 1276-1279.

Michalek, J.E., Wolfe, W.H., and Miner, J.C. (1990), "Health status of Air Force veterans occupationally exposed to herbicides in Vietnam. II. Mortality," Journal of the American Medical Association, 264, 1832-1836.

Michalek, J.E., Pirkle, J.L., Caudill, S.P., and Tripathi,R.C., Patterson Jr, D.G., and Needham, L.L. (1995)., "Pharmacokinetics of TCDD in veterans of Operation Ranch

Hand: 10-year follow-up," (To appear in the Journal of Toxicology and Environmental Health).

Palkovitz, M., Elekes, I., Lang, T. and Patthy, A. (1986). Taurine levels in discrete brain nuclei of rats J Neurochem 47, 1333-1335.

Sharma, H.S., Nyberg, F., Olsson, Y. and Dey, P.K. (1990). Alterationof substance P after trauma to the spinal cord: an experimental study in the rat. Neuroscience 38, 205-212.

Skilling, S.R., Smullin, D.H. and Larson A.A. (1990). Differential effects of C- and N-terminal substance P metabolites on the release of amino acid neurotransmitters from the spinal cord: potential role in nociception. Journal of Neuroscience 10, 1309-1318.

Smullin, D.H., Skilling, S.R. and Larson, A.A. (1990). Interactions between substance P, calcitonin gene-related peptide, taurine and excitatory amino acids in the spinal cord. Pain 42, 93-101.

Wolfe, FW., Smythe, H.A., Yunus, M.B., Bennett, R.M., Bombardier, C., Goldenberg, D.L., Tugwell, P., Campbell, S.M., Abeles, M., Clark, P., Fam, A.G., Farber, S.J., Fiechtner, J.J., Franklin,. C.M., Gatter, R.A., Hamaty, D., Lessard, J., Lichtbroun, A.S.,

Masi, A.T., McCain, G.A., Reynolds, W.J., Romano, T.J. Russell, I.J. and Sheon, R.P. (1990). The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Arthritis and Rheumatism 33, 160-172.

Wolfe, W.H., Michalek, J.E., Miner, J.C., Rahe, A., Silva, J., Thomas, W.F., Grubbs, W.D., Lustik, M.B., Karrison, T.G., Roegner, R.H., and Williams, D.E. (1990), "Health status of Air Force veterans occupationally exposed to herbicides in Vietnam. I. Physical health," Journal of the American Medical Association, 264, 1824-1831.

Wolfe, W.H., Michalek, J.E., Miner, J.C., Needham, L.L., and Patterson Jr, D.G. (1992), "Diabetes versus dioxin body burden in veterans of Operation Ranch Hand," in Organohalogen Compounds. Volume 10, Toxicology, Risk Assessment and Management, pp 279-282, Helsinki: Finnish Institute of Occupational Health.

6. APPENDIX

In our notation, Flanders et al Model 2 for subject i, i=1,2, ..., n, is given by

$$X_{2i}=a+bX_{1i}+cY_i+\varepsilon_i,$$

where X_2 is measured dioxin, X_1 is true dioxin, Y is the health variable and ε is error. The four variables, X_1, X_2, Y and ε are assumed jointly normally distributed with $E(X_1)=\mu_1$, $VAR(X_1)=\sigma_1^2$, $E(Y)=\mu_y$, $VAR(Y)=\sigma_y^2$, $E(\varepsilon_3)=0$ and $VAR(\varepsilon)=\sigma^2$ and X_1, Y and ε are assumed mutually independent. Thus health (Y) and measured dioxin X_2 are bivariate normal with mean vector $(\mu_y,\ a+b\mu_1+c\mu_y)$, $VAR(X_2)=c^2\sigma_y^2+b^2\sigma_1^2+\sigma_\varepsilon^2$, and $COV(Y,X_2)=c\sigma_y^2$. Therefore $E(Y|X_2=x_2)=\gamma_0+\gamma_1 x_2$, where

$$\gamma_0=\mu_y-\left\{\frac{c\sigma_y^2(a+b\mu_1+c\mu_y)}{c^2\sigma_y^2+b^2\sigma_1^2+\sigma_\varepsilon^2}\right\}$$

and

$$\gamma_1=\frac{c\sigma_y^2}{c^2\sigma_y^2+b^2\sigma_1^2+\sigma_\varepsilon^2}.$$

But Flanders et al also assume $E(Y|j=0)=E(Y|j=1)$, where j=0 for control subjects and j=1 for index subjects, which implies

$$\gamma_0+\gamma_1E(X_2|j=0)=\gamma_0+\gamma_1E(X_2|j=1).$$

But $E(X_2|j=1)>E(X_2|j=0)$. Hence, $\gamma_1 \equiv 0$ and, therefore, $c\equiv 0$. Thus, Model 2 and their assumption that $E(Y|j=1)=E(Y|j=0)$ imply that Y and X_2 are unrelated. As a result, $k\equiv 0$ in their equation (8) and so their conclusions regarding the check mark pattern and reverse causation are not correct.